I Sell Seashells by the Seashore and My Name is Jack:

A Critical Look at Pelham, Mirenberg, and Jones's (2002) Findings

Marcello Gallucci

Free University, Amsterdam,

The Netherlands

Authors' Notes: I wish to thank Jamie De Coster, Marco Perugini, Sander Koole, Charles Judd, Nathalie Wyer, Paul van Lange, and the reviewers of the article for their insightful help and comments. Correspondence concerning this article should be addressed to Marcello Gallucci, Department of Social Psychology, Free University, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. E-mail: m.gallucci@psy.vu.nl.

Abstract

According to a new hypothesis based on implicit egotism, people gravitate toward cities, states, and careers with names similar to their own names. To support this hypothesis, Pelham, Mirenberg, and Jones (JPSP, 2002, 82(4) 469-487) report a series of results regarding distributions of names in different cities, states, and jobs. In the present article new analyses of the original data are reported, showing that the hypothesis is not supported for the large majority of names considered by the authors, and for some names even the opposite result is found. In addition, a meta-analysis reveals that either the data are unreliable, or the hypothesis can not be supported in the whole population of names. Overall, the original data give no support of the idea that implicit egotism influences major life decisions.

The fascinating history of scientific enterprise is studded with hypotheses that have challenged our understanding of reality, the way we look at our world, and our comprehension of the causes underlying every-day life events. Among those challenging hypotheses, we can surely number a new social psychological hypothesis, recently proposed by Pelham, Mirenberg, and Jones (2002), which states that people prefer to live in places and to pursue careers with names similar to their own names. Behind this hypothesis is the notion that people like to feel good about themselves (Allport, 1961; Greenwald & Banaji, 1995), thus they like objects associated with the self (Beggan, 1992), including the letters of their names (Nuttin, 1985; 1987) or the numbers of their birthdays (Miller, Downs, & Prentice, 1998). Consequently, they prefer things (like cities, states, or jobs) that have associations with those objects. The consequence of this notion, the hypothesis goes, is that people named Louis prefer to live in Saint Louis and people called Florence choose to live in Florida.

To test this intriguing hypothesis, Pelham et al. conducted a large number of studies, reported a series of significant results, and concluded that the data support their hypothesis. The aim of my contribution is to challenge this conclusion. I propose that the original analyses can be questioned on the following ground: In all the studies reported by Pelham et al., the sampled units are not the individuals, as the authors' analyses implicitly assume, but rather the names, towns, and jobs the authors consider. The data structure in each of the original studies is a nested data structure, with individuals nested under names and towns (or jobs). Pelham et al.'s analyses seem to ignore this nesting structure and treat the individuals as units, casting doubt on the adequacy of the statistics that are used, the test of the hypothesis, and the generality of the effects.

In this contribution I have re-analyzed all of the data reported by Pelham et al. in the original article, and found little evidence in support of their hypothesis. Using statistical analyses that take into the account the structure of the data, I show that the hypothesis is not supported in the great majority of the tests, and in some cases its opposite is statistically verified. Specifically, the data allow reliable tests of the hypothesis for 74 names or numbers. Fifty-five of these 74 cases give no significant result whatsoever, four yield a significant result contrary to the hypothesis, and only 15 yield a significant result in support of the hypothesis[1]. For the remaining data which do not guarantee a reliable test of the hypothesis, I show that the method used by the authors can yield significant results due to spurious effects and sampling biases. Finally, treating all the names used by the authors as randomly sampled, I use Bernoulli's law of large numbers to show that, within the boundaries of my approach, either the hypothesis is not supported in the population, or that the data are likely to be biased. Overall, my analyses suggest that the data reported by the authors do not support their hypothesis.

The Original Data and Method

The original research deals with the hypothesis that people gravitate toward cities, states, or careers that have names similar to their personal names or birth dates. For the sake of brevity, in this contribution I will talk about names whenever I refer to objects related to self (names or birthdays), and to places whenever I refer to the target of decisions (cities, states, careers). Therefore names are sometimes birthdays, and places are sometimes cities, sometimes states, and sometimes careers, depending on the specific study under discussion. I will use the expression name-place match to refer to the combination of name of a person and name of a place that share the same initial letters (Louis-Saint Louis), whereas I will use the term name-place mismatch to refer to pairs of personal name and name of a place that do not share the same initial letters (Louis- Toronto).

The original empirical evidence is based on 10 studies using data collected from archival records publicly available on the internet. In each study, the authors chose a number of personal names and a number of places, and counted how many people with each name lived in places with names similar to the personal name. As a general method, the authors looked at the probability of occurrence of name-place matches, they estimate the expected probabilities due to chance, they compare the two probabilities, and draw conclusions concerning the hypothesis. Because the 10 studies reported by Pelham et al. differ in the number of names considered, in the way the expected probabilities are computed, and the way statistics should be applied, I will treat the studies in different sections, grouping the studies depending on their shared characteristics[2].

Overall, there are three groups of studies that share the same characteristics: a) Study 5, the "Saint cities study", where only name-place matches are considered and the expected probabilities of names are drawn from the entire population of American names; b) Studies 2, 3, 4, and 6, where a set of names (more than two) and a set of places (more than two) are cross-tabulated, providing data about both name-place matches and mismatches; c) Studies 1, 7, 9, and 10 where two name-place matches are compared in 2x2 tables.

As I presently show, those three categories of studies are not equivalent, for they have different degrees of reliability, different sampling biases, and require different considerations. The first category includes the best and most accurate data, the second category includes data which are less reliable but still workable, the third category represents poor sampling and share features of anecdotal data. I therefore start my analysis with the best data set available (the first category or Study 5 in the original article).

The Saint City Study

The rationale of Study 5 is the following. If people gravitate toward cities that remind them of their own name, we should observe that cities featuring a personal name (as Saint Louis, Saint Joseph, Saint Marie, etc.) should attract people named with that name (Louis, Joseph, Marie), more than what we would expect by chance. Upon this idea, the authors considered all the 35 American cities named after a Saint (eight with a female name, 27 with a male name), collected the actual frequencies of people with the Saint name living in the city, and compared those frequencies with the frequencies of each name in the whole American population (see Table 1 here and table 8 in the original contribution[3]).

To test the hypothesis, the authors used an overall test, which is based on the comparison between the sum of the actual probabilities of name-place matches (column 3 in Table 1) and the sum of the expected probabilities given by chance (column 2 in Table 1)[4]. Because this overall test ignores the fact that the units of analysis are names and states rather than individuals, this overall test does not test the hypothesis. The overall test tells us only if the two distributions (actual and expected probabilities) are different (Cramer, 1999). Obviously, this is necessary but not sufficient to test the hypothesis. This overall test, in fact, might be significant even if only one name-place match out of 27 is significantly different from chance, while the rest of the name-place matches are as likely as chance or even less likely than chance.

The correct test of the hypothesis should generalize the effect across names. We therefore need to test how many names reveal a significant effect in support of the hypothesis, how many are not in support of the hypothesis and how many, if any, are against the hypothesis (i.e., significantly less than chance). Only by considering this consistency of the effect across different names we can evaluate if the hypothesis is supported by the data. Table 1 reports the tests and the significant effects I found for the "Saint" data set. As regards female names, I observe that the expected and actual name frequencies are different ($\chi^2(7)=69.0$ $\underline{p} < .01$), but only two names (Mary and Clair) are statistically significant from chance. This case exemplifies the aforementioned inadequacy of the overall test. For female names, in fact, the overall test ($\chi^2=69.0$) is significant in that one name-place match (Mary

χ^2=59.89) accounts for almost all the differences between the expected and the actual frequency distribution.

As regards male names, the expected and observed frequency distributions are also different ($\chi^2(26)=606.9$ $p < .01$), and 10 individual name-place matches are significantly different from chance. Of these 10 significant results, however, six are in support of the hypothesis and four are against the hypothesis, that is people with a name matching the city name are significantly less likely to live in that city. For male names, the inadequacy of the overall test is even more pronounced than for female names: In Saint Louis the observed frequency of the name Louis is 2266 and the expected frequency is 1495, clearly in support of the original hypothesis. But if this one city is left out and the data for all the individuals for the 26 cities are used, the observed frequency of name matches is 1729 and the expected frequency is 1981, clearly in the opposite direction from the hypothesis. So, the result reported in the original article, ignoring cities and names as the unit, derives simply from a single city-name matching[5].

All things considering, the best data set available from the original research yields results that can hardly be considered supportive of the name-place matching hypothesis. Out of 35 names, for 23 there is no evidence in support of the hypothesis, for eight the hypothesis is supported, and for four the opposite of the hypothesis is found[6]. Disregarding test of significance, the analyses for all 35 names reveal a difference between the observed and the expected frequency in the predicted direction for 16 names, whereas for 19 names, this difference is in the wrong direction.

Such small probability of significant effects should also be evaluated in light of the sampling effects that might exist. I will discuss the meaning of those results in light of potential sampling errors in a more detailed and statistical fashion later on. I now turn to the second type of studies Pelham et al. described in their article, and show that the situation is not much different from the one above.

The Cross-tabulation Studies

Studies 2,3,4, and 6 are conducted as follows: N names and N states (or cities) are selected, such that each name (e.g., Florence, Georgia, Louise, and Virginia) can be matched with a state (e.g., Florida, Georgia, Louisiana, Virginia) according to its initial letters. The frequency of people living in each state with each of the names are collected from public archives, and an NxN table is constructed. To test their hypothesis, the authors computed an overall test (with 1 degree of freedom), based on the difference between observed and the expected probabilities of the name-place matches. All of the studies, according to the authors, yield significant results in support of the hypothesis.

As for the Saint city study, an overall test is not informative in that one name can be responsible for the overall effect, as for Saint Louis. In contrast to the Saint city study, however, we can not simply test each name-place match in order to count the number of significant results, because we do not have the expected distributions in the whole population. Thus, the expected frequencies (due to chance) can only be computed using the marginal frequencies of the tables. Because the tables include very few names and cities (from four to eight), those data pose several different problems: First, the population-wise distribution of a name is computed out of four, seven, or eight cities, rather than across the thousands of cities that form the real population. Thus, substituting only one city in a table may be enough to change all the results. Second, the expected frequency of a given name in the letter-matching place is not computed based on the population of other names in that place, but rather on the few particular names selected in the study. Thus, changing only one name in the table can yield completely different results. Third, the effects one might find can be due to the proportion of people in the mismatching cities rather than in the matching city (see below for examples).

To avoid arbitrary effects, we should test the hypothesis on the NxN tables using a logic that considers names as units of analysis, across which generalization is sought. Consider a case with four names and four cities. The hypothesis states that if my name is A, it should be more likely that I live in Apolis than in Bpolis, Cpolis and Dpolis[7]. Thus, the first necessary condition is that the distribution of probabilities of people named A in the four cities should be different from the expected (from chance) distribution. The second condition is that the probability of A in Apolis should be higher than chance. The third condition is that the probability of A in all the other cities excluding Apolis should be random (equal to the expected). The necessity of those three conditions can be easily appreciated in Figure 1. Figure 1a shows the probabilities provided in Pelham's et al. Study 3, of living in eight

Canadian cities if your name is Tor* (read names beginning with Tor). The values in the figures are the differences between the expected probability and the actual probability of each cell, such that 0 indicates chance probability, positive values mean more likely than chance, and negative values mean less likely than chance. I chose this example in that for the name Tor* , the hypothesis is clearly verified, so we can appreciate how the method I propose applies to a positive case.

As can been seen in Figure 1a, the probability of the match Tor-Toronto is higher than chance, and the rest of the probabilities are uniformly low. This means that, as the hypothesis predicts, people named Tor* are more likely to be found in Toronto than they are in the other seven cities. This means also that if we test the distribution of Tor* in the eight cities against the expected distribution, we find a significant effect, whereas if we remove Toronto, we find no significant effect. The first test tells us that people named Tor* do not distribute randomly in the eight cities, the second test tells us that this is due to the fact that people named Tor* are over-represented in Toronto.

Why do we need the second test? The reason is that the expected frequencies are dependent on the frequencies of the few names in the table, thus without the second test we can not attribute a significant effect to the name-place match. To appreciate this, consider the probabilities of living in the eight Canadian cities if you are called Edm*, depicted in Figure 1b. Here also we see that the match Edm-Edmonton is more likely than chance, but in this case the effect is due to the fact that people named Edm* are under-represented in Toronto and Hamilton, with a consequent increase in the remaining cities. As a consequence, it seems that people called Edm* gravitate toward Edmonton, whereas they actually gravitate, with higher probability, toward Calgary and Vancouver, and, with around the same probability, toward London and Winnipeg. Thus, for the name Edm*, the effect is not due to the letter matching, but instead to the distribution of people in Toronto, Hamilton, and Calgary, clearly outside the reach of the hypothesis.

The previous reasoning leads to the following test: a) Compute the Chi-square test for the distribution of people named X in the N cities (against the expected values); b) compute the Chi-square test for the distribution of people named X in the N-1 cities, excluding the matching city (against the expected values re-computed after deleting the matching city). Because the requirement of randomness in the second distribution may seem too conservative, we should at least test that the distribution of people, excluding the matching city, should be more similar to the expected distribution than the distribution including the matching city. This can be accomplished by computing the ratio of the two Chi-squares (with and without the matching city). The ratio of two Chi-squares, each divided by its degrees of freedom, distributes as a F (cf. Isaac, 1999), and its significance can be evaluated as in the ANOVA model, using the F distribution[8]. If the ratio is significant, the hypothesis is supported for that particular name. Then we can count how many names support the hypothesis.

Using this method I have re-analyzed all the tables cited by the authors in Studies 2,3,4, and 6 (see Tables 2 and 3 in this contribution). Study 2 considered one table of eight names (table 2 in the original article). I obtained one significant result. Study 3 considered one table of eight names (table 3 in the original article). I obtained one significant result. Study 4 considered four tables of four names each (tables 4,5,6, and 7 in the original article). I obtained two significant results for table 4, one significant result for table 5, two significant results for table 6, and one significant result for table 7. Study 6 considered one table of seven birthdays, cross-tabulated with seven cities featuring names which include numbers (Two Harbours, Three Oaks, etc.). Here I found no significant result (see Table 4). In this last study, the distribution of people with each birthday across the seven cities is equal to the expected distribution, so no further test is necessary to reject the hypothesis.

To test the validity of my method, it is worth noting that it yields a significant result for Tor-Toronto, whereas it yields a non significant result for Edm-Edmonton, as we would expect by looking at the graphs in Figure 1. To reinforce its validity, note also that my method yields a significant result for Texa-Texas, and a non significant result for Illi-Illinois, an outcome that we would expect after inspecting the probabilities depicted in Figure 2.

Overall, those studies allow us to test the hypothesis 39 times. I obtained eight significant results in support of the hypothesis. This means that across those studies, 79% of the names yield no effect whatsoever. It seems doubtful that these represent support for the hypothesis when the hypothesis is refuted 79% of the cases[9]. Furthermore, this 21 % of positive cases should be evaluated in view of nonrandom sampling of names used by the authors, and the very small number of names employed to compute the probabilities (see the meta-analytic section below).

2x2 Table Studies

Four studies of the original research are based on 2x2 tables. In Study 1 two names and two matching cities are selected, the frequencies of people with those names in those cities are cross-tabulated, and the frequencies are evaluated with a Chi-square. In Studies 7,9, and 10, the same operation is conducted for jobs instead of cities.

For these studies, of course, no new analysis can show different results. Are those data therefore reliable tests of the hypothesis? I propose that they are not, for a simple reason: Because the unit of the analysis should be names and not people, each 2x2 table should be treated as hardly more than a single observation. As with any test performed on an extremely small sample size, serious problems soon arise: First, the fact that we find a significant result in a 2x2 table does not imply that the effect will generalize if we enlarge the number of names in the sample. Second, 2x2 tables can yield a significant effect for an infinite number of reasons beyond the correspondence between names and places. I now provide evidence relevant to these two major problems.

Assume that 2x2 tables were reliable ways to test the hypothesis. If so, I can test the hypothesis in the 2x2 tables contained, for instance, in the 8x8 table of Study 2 or Study 3. In each NxN table, in fact, there are (N-1)N/2 distinct 2x2 tables which cross-tabulate two name-city matches. I can therefore use these data, which according to the authors provide evidence in support of the hypothesis, to construct 28 2x2 tables that could be examined individually. For each data set (original tables 2 and 3), the 28 tables can be formed matching, in turn, every two names with the corresponding two cities (that is, the first table cross-tabulates Tor and Vanc, with Toronto and Vancouver; the second Tor and Ott with Toronto and Ottawa, and so on up to the last table which cross-tabulates Ham and Lo with Hamilton and London). For each table we can compute the Chi-square according to Pelham et al. method.

The results I found clearly show the inconsistency of results yielded by 2x2 tables. As regards the data of Study 2, out of 28 tests, I found 17 significant results, 14 in favor of the hypothesis and three against the hypothesis. As regards to the data of Study 3, out of 28 tests, I obtained 11 significant results in support of the hypothesis, and one against the hypothesis. Thus, less than half of the tests show a significant effect and half a non-significant effect. Consequently, the fact that one 2x2 table yields a significant effect for a name, does not imply that the name would show a consistent effect when compared with other names. This means also that when a name does not show a consistent effect when compared with many names, it can still produce spurious effects when compared with only one name. We can find evidence of this bias by analyzing the birthday study (Pelham et al.'s table 7). If we use a 2x2 approach on table 7, we obtain four significant results in favor of the hypothesis, despite that fact that table 7 is not different from a random table, as indicated by my analysis of NxN tables and by a standard Chi-square test ($\chi^2(36)=42.95$, p=.19).

Do these analyses using 2x2 tables only fall short in generalizing the hypothesized effect? Unfortunately not. The major problem we face by using 2x2 tables is that the sample of names one uses is so small as compared with the population of names, that the effects can be due to a variety of spurious factors affecting the data. To show the danger associated with 2x2 tables, I have taken the original table 2, and explicitly remove the possibility that any effect is due to the name-letter matching. This can be done by exchanging column 1 with column 8, column 2 with 1, column 3 with 2, and so on, with the resulting effect that the hypothesis becomes supported when Georgi-California, Cali-Texas, Texa-Florida, Flori-Illinois, Illi-Pennsylvania, Penny-Ohio, Ohi-Michigan, Michi-Georgia are more likely than other pairs. I then constructed the 28 2x2 tables again, and ran my tests. Out of 28 tests, I obtained 15 significant results, 12 in favor of the hypothesis and three against the hypothesis. Consequently, even when I removed the possibility that the statistical effects are due to the matching of names and places letters, I was able to produce about the same proportion of significant results than by matching names and cities according to the letters in their names. And this is not all.

We can take this reasoning a step further. In a 8x8 table there are 40,320 possible orders of the columns, in 14,833 of which no name is matched with the corresponding original place (cf. Roberts, 1984). Those 14,833 orders produce 415,324 2x2 tables (14,833 x 28) in which names and places are mismatched. Importantly, because in these 415,324 tables the correspondence between names and places is removed, almost no effect in the original direction should be found (more exactly, no more than 5%), both under the hypothesis that a name-letter effect does exist, and under the null hypothesis of no effect at all. That is, as along as no spurious effect occurs, almost no effect should be found in these tables. Instead, using the data in table 2, I found 94,887 2x2 tables (22%) yielding a significant result in direction of the hypothesis (63,734 or 15% for table 3). Thus, even when non effect should be found irrespective of the null hypothesis we intend to test, we still observe a substantial number of statistical effects that are neither random nor justifiable with the authors' hypothesis. Consequently, they are spurious effects due to the fact that a 2x2 tables samples a very tiny part of the existing population of names and places. On the basis of these considerations, we can reasonably ask whether single 2x2 tables presented in Pelham et al. Studies 1,7,9, and 10 yield a real effect or are drawn from the same population of our 94,887 tables producing spurious effects. The answer is that we simply do not know. This is another reason why 2x2 tables are not suitable for testing the name-place matching hypothesis[10].

Meta-Analysis: Is There a Sampling Problem?

Throughout this contribution, I have argued that in light of the potential sampling problem of the studies conducted by Pelham et al., even the few significant results I obtained can be used to show that the hypothesis is not supported. In this section I provide evidence in favor of my argument.

It is clear that the sampling problem of the original contribution does not concern the number of people involved in the studies (which is large), nor the procedure employed to sample them (which is random). The sampling problem in the original studies concerns the number of names (self-related objects) and cities (decision-targets) considered in each study. Specifically, the number of cities and names used to compare the probabilities of the target names is so small that we cannot confidently draw reliable conclusions from these studies. If we could, we would conclude that the hypothesis is not supported. That is, I now show that if the original studies are unbiased, they indicate that the hypothesis is not supported in the whole population. Conversely, if one still sustains that the hypothesis is supported in the population, than the studies must be necessarily biased and therefore inconclusive.

How can I prove my point? By employing Bernoulli's well-known law of large numbers (cf. Isaac, 1995), we can determine, a priori, the relationship between a) the number of sampled units (i.e., the number of names in a study) and b) the probability of obtaining a supporting result. Fortunately for my case here, the relationship between number of sampled units and probability of a supporting result is different, depending on whether the hypothesis we are testing is verified in the population (i.e., supported if all possible members of the population are considered) or falsified in the population (i.e., not supported if all possible members of the population are considered).

Let me clarify the nature of this relationship using an example. Assume I want to test the hypothesis that Jack is taller than the average American person. Assume further that Jack is indeed taller than average, thus my hypothesis is verified in the population. In the first experiment, I compare Jack with one person I randomly choose from a given population, in my second experiment I compare Jack with 10 people, in my third with hundred people and so on. In the first experiment, I might find someone who is taller than Jack or smaller than him, and thus find some non supporting result, more or less depending on my luck. Also in my second experiment I might find some group that is on average taller than Jack, but the chance of finding such a group is smaller than the chance to find a single person taller than Jack. Bernoulli's law of large numbers assures us that the larger is the sample of cases I compare Jack with, the more the average of the sample will be similar to the average of the population, and thus the higher the probability of finding supporting results. Thus, if the hypothesis is verified in the population, the larger the number of sampled units, the higher the probability to verify the hypothesis.

Assume now that Jack is shorter than the average American person, so my hypothesis is falsified in the population. In my first experiment, I might very well find a person taller than Jack, but I might also select a person smaller than him. In larger samples, however, the probability of finding a comparison group with an average height lower than Jack's decreases, because Jack is indeed smaller than the population average. Thus, if the hypothesis is falsified in the population, the larger is the sample of units we employed, the lower is the probability of verifying the hypothesis[11].

To summarize, from the foundations of statistical sampling we know that if a hypothesis is supported in the population, the correlation between the probability of supporting results and number of sampled units is positive, whereas if a hypothesis is not supported in the population, the correlation is negative. Pelham et al.'s studies allow us to compute such a correlation, because the authors provide 15 databases of different numbers of names. In (original contribution) tables 1a, 1b, 10a, 10b, 12 and 13, two names are compared, so the sampled units are two; in tables 4, 5, 6, and 7, the sampled units are four, in tables 2, 3, and 8a the units are eight, in table 8b there are 27 units, and in table 9 there are seven units. Table 5 in this contribution summarizes these numbers: For each original table the number of names involved in the test and the probability of significant supporting result is reported, upon which the correlation is computed.

The correlation between the number of names in the study and proportion of supporting results is -.54, which is negative, high and significantly different from zero, even with only 15 observations. We therefore observe a strong negative association between the number of sampled units tested in the studies and the proportion of significant results one obtains in support of the hypothesis. As we know, this is the signature of hypotheses that are not supported in the population. In fact, if we accept the data as unbiased, we should expect that by enlarging the sample of names, the probability of supporting results would decrease, vanishing as we approach a perfect sampling. If the hypothesis is supported in the population despite this negative correlation, we could explain this result only as the effect of a fundamental fault in the data, which makes studies with more names less and less reliable.

Thus, by simply looking at the number of significant results and the number of names involved in each study, we can conclude that the hypothesis and the data cannot be both valid at the same time. Sustaining otherwise would be as paradoxical as saying that the height of one person randomly chosen in a population is a better estimate of the population mean that the mean computed across hundred persons[12].

An important characteristic of the previous meta-analysis it that it helps us understand the status of the hypothesis even beyond and independently of the tests I have conducted on the original data. In the previous meta-analysis, in fact, I computed the probability of significant results according to the results I obtained. It is however possible to conduct the same meta-analysis using the statistics reported by Pelham et al.'s original article. According to the authors, each data set can be considered in favor of the hypothesis because their analyses yielded significant results. Thus, the probability of supporting results can be obtained by simply counting, in each data set, how many names show an actual frequency higher than the expected frequency. Figure 3 presents the probability of supporting results according to the authors' analyses as a function of the number of names considered in each data set. The correlation between number of names and supporting results is now -.74, which is even more negative than in the previous analyses. Figure 3 clearly suggests that the probability of finding a supporting result tends toward zero as we increase the number of names in the sample. Thus, also using Pelham et al.'s statistical approach, the original results presented by the authors are still not in favor of the hypothesis[13].

Conclusions

Pelham et al. have carried out a very intriguing line of research guided by a new and original hypothesis: Major life decisions are influenced by unrelated items that assume a new meaning when associated with the self. Their effort should be applauded for it is rare to find such enthusiasm in applying social psychological concepts to real life decisions, decisions that are undoubtedly more important than letters on a computer screen. Pelham et al.'s hypothesis, furthermore, might have challenged our idea that such important decisions as moving to a city or choosing a career are influenced not only by future prospects, salaries, opportunities for the family, migration flows, economical trends, potential improvements of life conditions, and realization of personal dreams, but also by supposedly unrelated things, like letters in a name. Strangely enough, this is one of the fundamental aims of science, to challenge common sense and pre-existing views of reality.

Another aim of science is to promote hypotheses with compelling empirical evidence. The empirical findings, I have argued, do not support the name-place matching hypothesis. I have built my argument as follows: First, I have accepted the hypothesis as derived by the original authors, without changing it or criticizing their theoretical apparatus. Second, I have assumed that the data reported by the original authors were sufficient to verify the hypothesis. Thus, starting from the same theory and the same data, I have approached the hypothesis with a series of statistics which are, in my opinion, more appropriate than the original ones. I have also tried to attribute the statistical effects I found to their likely causes, avoiding the confusion between undistinguished significant effects and significant effects supporting the hypothesis. Finally, I have generalized the results in order to identify the effects of sampling biases and limitations of the empirical data, using consolidated statistical theory and logical arguments.

My analyses suggest that the original data provide little evidence in support of the hypothesis. As compared with the original authors' analyses, I found that very few name-city pairs are more likely than chance, some pairs are even less likely than chance, and the majority are as likely as chance. I also found that there is a strong association between the number of names in the sample and the probability of finding a significant result, suggesting that if the data are reliable, we can conclude that the hypothesis is not supported in the population.

In conclusion, my analyses strongly suggest that implicit egotism does not influence major life decisions such as moving to a city or pursuing a career. Of course, new data, better statistics, and more accurate sampling may challenge my conclusions as well. After all, the first and most important aim of my contribution is to stimulate supporters of this hypothesis to provide more robust evidence and more careful analyses in support of it. The second aim of mine, fortunately, is to assure my dear friend Jack Priston that he is not doomed to end up in jail and be a prisoner for the rest of his life.

Footnotes

1 In this contribution significance is evaluated at $\alpha = .05$, even when multiple tests would require a lower α (Bonferroni's correction). I made this choice in favor of the authors' hypothesis and because a lower α would decrease the power of the tests, an issue that may result important when the null hypothesis is not rejected. When Bonferroni's correction is considered, out of 74 cases I obtained 61 (82%) non significant results, 3 (4%) significant results contrary to the hypothesis, and 10 (13%) in support of the hypothesis.

2 I will consider 9 out of 10 studies because the data of Study 8 are not reported in the original article.

3 To avoid confusion, I refer to tables in Pelham et al.'s as tables in the original article. In addition, I do not capitalize the word "table" when it refers to a table in the original article, and I capitalize the word when it refers to a table in the present article.

4 Although the authors did not specify which test they used, I was able to infer it (a squared z-test on proportions that distributes χ^2), thanks to the careful calculations of an unanimous reviewer of this paper. I am grateful for this help. Note that this method of testing the difference between the expected and the observed distribution is highly questionable (Cramer, 1999), because it is based on 1 degree of freedom rather than k-1 degrees of freedom (where k is the number of names in the sample). I do not discuss this problem in the text, because the overall test is not appropriate to test the hypothesis even if it is conducted with a less questionable approach, such as a Chi-square test with k-1 degrees of freedom.

5 I wish to thank Charles Judd for suggesting this intriguing result. Note that this result is not affected by the statistical power of the tests I have conducted, so it can help to rule out possible concerns related with this issue.

6 A possible limitation of the analysis conducted on single names in Table 1 (original table

8), is that the expected frequencies for some of the names are very small (<5). To overcome this limitation, one can analyze only the names that yield expected frequencies greater than 5 (reported in boldface type in the original table 8). Considering only those names, one finds that the percentage of supporting results goes from the original 22% to 33%, but also the percentage of results against the hypothesis increases from the original 11% to 26%. Thus, the fact that some names yield small expected frequencies does not influence the conclusions of the analysis.

7 Note that, in favor of the authors' case, I am not considering the "priming effect", that is parents that name their children after the state or the city the children are born in (people in Rome are often called Romolo because Romolo was the founder of Rome). Considering this effect, however, yields very interesting results: In Study 3 data about residents of four states and data about immigrants in those states are presented. From those data can be evinced that, for every name and across names, non-immigrants (total minus immigrants) are more likely than immigrants to live in a matching state. For instance, for female names (tables 4 and 6), the difference between observed and expected probabilities of name-place matches is .06 for immigrants, and it is .12 for non-immigrants. Because the effect of implicit egotism is ruled out by design in the non-immigrant sample, we can consider non-immigrants as a control group. Thus, I found that the hypothesized effect is stronger in the control group than in the experimental group (immigrants), another result in the opposite direction of the hypothesis.

8 More precisely, the ratio of two chi-squares, χ_1^2 with n degrees of freedom, and χ_2^2 with m degrees of freedom, each divided by its degrees of freedom, distributes as F with n and m degrees of freedom. That is, $F_{n,m}=(\chi_1^2/\chi_2^2)(m/n)$. In our particular case, the degrees of freedom are obtained subtracting 1 from the number of places we consider. For 8 cities, for instance, the chi-square with the matching city has 7 degrees of freedom, whereas the chi-square without the matching city has 6 degrees of freedom. The F ratio is than obtained as

$F_{7,6}=(\chi^2_{with}/\chi^2_{without})(6/7)$.

9 It should be noted that the analyses I have conducted on the NxN tables are by no means conservative tests. A more stringent test of the hypothesis should also test the distribution of names within one city. In fact, the hypothesis not only implies that people should prefer places with their names more than places with different names (an effect column-wise), but also implies that places should attract more people with a similar name than people with a different name (an effect row-wise). When the hypothesis is tested also row-wise, out of the 39 places, four places show a significant result in favor of the hypothesis, and one against the hypothesis. Thus, a more stringent test of the hypothesis reduces the proportion of significant results to 10%. I gratefully thank Marco Perugini for suggesting this argument.

10 Given the unusual nature of the previous analysis, some detail is in order. First, for the sake of completeness, is worth mentioning that I obtained 28% of significant results in the opposite direction of the hypothesis (27% for table 3). Note that if the original data were in support of the hypothesis, this percentage should have been around 50%, because 50% of the mismatching tables are exactly the opposite of a matching tables. Thus, independently of what direction our hypothesis goes, we obtain an error rate of about 22%. Second, the careful reader may notice that in the 415,324 2x2 tables discussed in the text, every table appears more than once. This is not a problem because those repetitions are independent of the significance of the test performed on the table. For completeness, however, I have also performed the tests on the 1204 unique 2x2 tables produced by any possible orders of columns (with name-place mismatching), and I found 270 tables (22%) significantly in the direction of the hypothesis, corresponding to the same proportion of significant results I found for the 415,324 tables (15% for table 3). Third, it is important to note that the previous analysis is not meant to draw conclusions regarding tables 2 or 3, already discussed in the section about cross-tabulation studies. If we were to use the analysis for this purpose, in fact,

we would draw contradictory conclusions. On the one hand, we can test the null hypothesis that the original order in table 2 (table 3) is equivalent to any other mismatching order. Because only 2% (6%) of the mismatched orders produce an equal or greater number of significant results than the original order, we can reject this null hypothesis of equivalence with $p=0.02$ ($p=0.06$). This result can be surely interpreted in favor of the authors' case. On the other hand, we can test the null hypothesis that the effects are due to the name-place correspondence, by assessing the probability of obtaining no significant results when the correspondence is removed. Because this probability is 0.009 (0.003), we can reject the null hypothesis that the effects are due to the name-place correspondence with $p=0.009$ ($p=0.003$). Taken together, these results can be interpret either against or in favor of the authors' hypothesis, depending on one's viewpoint. Either way, these results do not contradict the analysis conducted on tables 2 and 3 in the cross-tabulation section, and certainly do not invalidate the evidence regarding the spurious effect produced by 2x2 tables. For readers interested in performing this kind of analyses, ad hoc SAS macros are available on request from the author. Without a fast computer, however, running these analyses requires a lot of patience.

11 The third case is that in all the experiments the probability is equal to the probability in the population, which entails a zero correlation but requires that all the experiments show the same proportion of significant results, which is not the case here.

12 It is important to note that the previous analysis is quite robust even if it is based only on 15 data sets. It is in fact very difficult to change the negative correlation I found into a positive one, without decreasing the proportion of significant results. As an example, given the present data, the correlation would become positive if we add a new data set with 50 names showing 100% of significant results, or four data sets of 20 names each showing 100% of significant results.

13 Note that the meta-analysis conducted on the original authors' results is not based on the acceptance of the statistical null hypothesis of single tests, so it is not influenced by the statistical power of those tests.

References

Allport, G.W. (1961). Pattern and growth in personality. New York: Holt, Rinehart, & Winston.

Beggan, J.K. (1992). On the social nature of nonsocial perception: The mere ownership effect. Journal of Personality and Social Psychology, 62, 229-237.

Cramer, H. (1999). Mathematical Methods of Statistics. NJ: Princeton University Press.

Greenwald, A., G. Banaji M.,R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102, 4-27.

Isaac, R. (1995). The pleasure of probability. New York: Springer-Verlag.

Miller, D.T., Downs, J.S., & Prentice, D.A. (1998). Minimal conditions for the creation of a unit relationship: The social bond between birthday mates. European Journal of Social Psychology, 28, 475-481.

Nuttin, J.M. (1985). Narcisism beyond Gestalt and awareness: The name letter effect. European Journal of Social Psychology, 15, 353-361.

Nuttin, J.M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. European Journal of Social Psychology, 17, 381-402.

Pelham, B. W., Mirenberg, M.C., and Jones, J.T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. Journal of Personality and Social Psychology, 82, 469-487.

Roberts, F. S. (1984). Applied Combinatorics. Englewood Cliffs, NJ: Prentice-Hall.

Table 1

Distribution of names in the Saint cities and corresponding test of the hypothesis (Study 5)

Name	Proportion		City	χ^2	p.
	Expected	Actual	Population		
Agatha	.000091	.000000	183	.02	.90
Anne	.001305	.000000	1703	2.22	.14
Bernice	.001381	.000000	133	.18	.67
Clair	**.000155**	**.000315**	**25376**	**4.21**	**.04**
Helen	.009068	.010676	1405	.40	.53
Marie	.004591	.005975	5021	2.09	.15
Mary	**.022972**	**.033901**	**11504**	**59.82**	**<.01**
Rose	.004141	.003431	583	.07	.79
Anthony	.002508	.003858	1296	.94	.33
Augustine	.000084	.000000	13057	1.10	.29
Bernard	.001523	.001600	1250	<.01	.94
Charles	.014408	.015509	21343	1.80	.18
David(s)	**.004549**	**.002035**	**2948**	**4.57**	**.03**
Elmo	.000126	.000000	1083	.14	.71
Francis	**.002432**	**.004752**	**2315**	**4.17**	**.04**
Gabriel	.000148	.000000	276	.04	.84
George	.014347	.012532	6942	1.59	.21
Henry	**.006720**	**.033755**	**474**	**56.33**	**<.01**
Ignance	.000007	.000000	1328	.01	.92
Jacob	**.001111**	**.005319**	**376**	**5.99**	**.01**
James	**.020204**	**.015049**	**10499**	**13.75**	**<.01**
Joe	**.002471**	**.005117**	**2345**	**6.64**	**.01**
John	**.029861**	**.022749**	**5187**	**8.83**	**<.01**
Joseph	**.013665**	**.008143**	**36349**	**81.29**	**<.01**
Leonard	.002038	.002132	469	<.01	.96
Louis	**.004168**	**.006206**	**358699**	**397.62**	**<.01**

Table 1 (continued)

Name	Proportions		City	χ^2	p.
	Expected	Actual	Population		
Mark(s)	.000679	.000000	113	.08	.78
Martin	.001477	.000000	77	.11	.74
Matthew(s)	.000536	.001037	1928	.90	.34
Michael	**.003717**	**.013210**	**757**	**16.33**	**<.01**
Paul	.005469	.005445	119736	.01	.91
Peter	.002414	.002956	2706	.33	.57
Stephen(s)	.001221	.000549	1823	.67	.41
Thomas	.007796	.013746	873	3.57	.06
Vincent	.001080	.000000	56	.06	.81

Note. Names yielding a significant result ($\alpha = .05$) are in boldface type. Names yielding a significant result ($\alpha = .05$) opposite to the hypothesis are underlined.

Table 2

Test of the hypothesis for names in Study 2 and 3

Name	χ^2 with Match	p.	χ^2 without Match	p.	F Ratio	p.
Cali	132.59	<.01	89.32	<.01	1.27	.39
Texa	**121.8**	**<.01**	**8.46**	**.21**	**12.34**	**<.01**
Flori	91.86	<.01	51.19	<.01	1.54	.31
Illi	108.54	<.01	108.92	<.01	.85	.59
Penny	395.33	<.01	381.44	<.01	.89	.57
Ohi	69.69	<.01	56.64	<.01	1.05	.48
Michi	95.34	<.01	52.93	<.01	1.54	.31
Georgi	41.87	<.01	41.3	<.01	.87	.58
Tor	**149.39**	**<.01**	**24.48**	**<.01**	**5.23**	**.03**
Vanc	16.91	.02	16.9	.01	.86	.58
Ott	43.07	<.01	43.13	<.01	.86	.58
Edm	12.92	.07	12.15	.06	.91	.55
Cal	84.09	<.01	51.43	<.01	1.40	.35
Win	51.83	<.01	42.5	<.01	1.05	.49
Ham	70.78	<.01	49.95	<.01	1.21	.41
Lon	37.94	<.01	19.48	<.01	1.67	.27

Note. The second and third column report the Chi-square test and significance concerning the distribution of people with the corresponding name in the four states considered. The fourth and the fifth report the test excluding the matching state. The sixth and seventh column report the Chi-square ratio and level of significance. Names yielding a significant result (α = .05) in support of the hypothesis are in boldface type.

Table 3

Test of the hypothesis for names in Study 4

Name	χ^2 with Match	p.	χ^2 without Match	p.	F Ratio	p.
Florence	**2374.30**	**<.01**	**175.24**	**<.01**	**9.03**	**.05**
Georgia	1434.40	<.01	137.03	<.01	6.68	.07
Louise	1031.79	<.01	602.69	<.01	1.14	.43
Virginia	**2094.59**	**<.01**	**151.45**	**<.01**	**9.22**	**.05**
George	806.63	<.01	614.71	<.01	.87	.50
Kenneth	319.05	<.01	161.37	<.01	1.32	.39
Louis	**4779.23**	**<.01**	**96.74**	**<.01**	**32.94**	**.01**
Virgil	1338.61	<.01	865.24	<.01	1.03	.46
Florence	**596.33**	**<.01**	**2.36**	**.31**	**168.80**	**<.01**
Georgia	323.13	<.01	181.15	<.01	1.19	.42
Louise	165.83	<.01	139.41	<.01	.79	.53
Virginia	**170.74**	**<.01**	**4.59**	**.10**	**24.82**	**<.01**
George	30.01	<.01	28.21	<.01	.71	.56
Kenneth	3.79	.29	3.66	.16	.69	.57
Louis	**126.85**	**<.01**	**2.09**	**.35**	**40.42**	**.01**
Virgil	158.73	<.01	77.35	<.01	1.37	.38

Note. The second and third column report the Chi-square test and significance concerning the distribution of people with the corresponding name in the four states considered. The fourth and the fifth report the test excluding the matching state. The sixth and seventh column report the Chi-square ratio and level of significance. Names yielding a significant result (α = .05) in support of the hypothesis are in boldface type.

Table 4

Test of the hypothesis in Study 6

Birthday	χ^2 with Match	p.
2.2	4.21	.65
3.3	11.75	.07
4.4	4.55	.60
5.5	5.19	.52
6.6	6.87	.33
7.7	4.97	.55
8.8	5.42	.49

Note. The second and third columns report the Chi-square test and significance concerning the distribution of people with the corresponding birthday in the seven cities considered. No other test is reported because the first necessary condition is already not satisfied, that is, for each birthday, the distribution of people in the seven cities is equal to the distribution expected from chance.

Table 5

Proportion of significant results supporting the hypothesis as a function of number of names considered in the study

Original Table	Number of names in the study	Proportion of significant results
1	2	1
1	2	1
2	8	.125
3	8	.125
4	4	.500
5	4	.250
6	4	.500
7	4	.250
8	8	.250
8	27	.220
9	7	1
10	2	1
12	2	1
13	2	1
14	2	1

Figure Caption

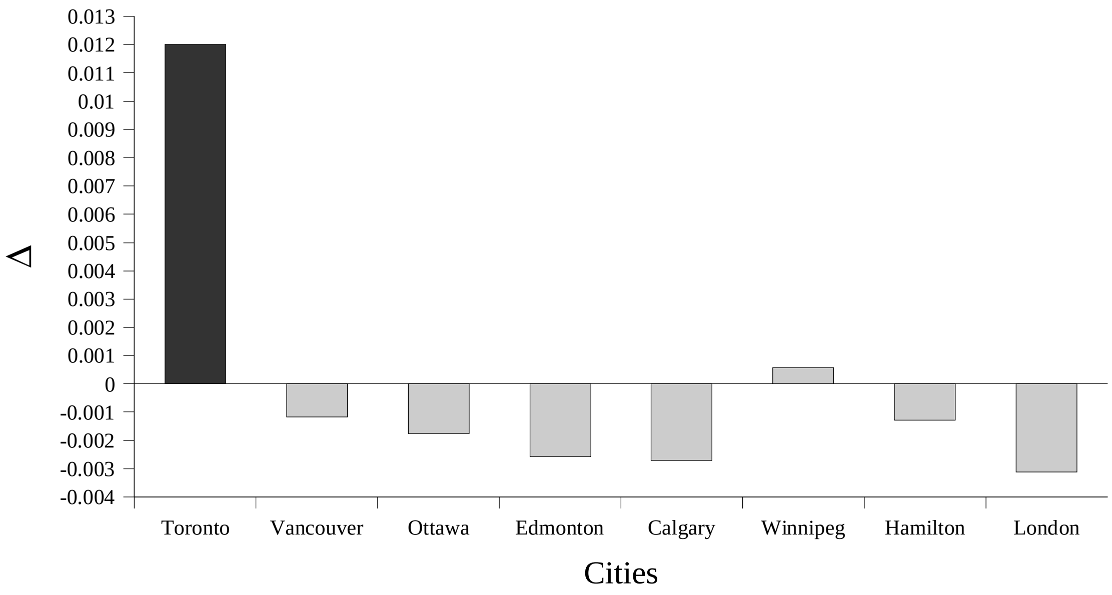
Figure 1. Differences between actual and expected probabilities of living in eight Canadian cities given a name. a) Names beginning with Tor. b) Names beginning with Edm. Δ indicates the observed probability minus expected probability, where probability is the frequency of a cell divided by the sample size. The probability of the name-place match is emphasized with a darker bar.

Figure 2. Differences between actual and expected probabilities of living in eight American states given a name. a) Names beginning with Texa. b) Names beginning with Illi. Δ indicates the observed probability minus expected probability, where probability is the frequency of a cell divided by the sample size. The probability of the name-place match is emphasized with a darker bar.

Figure 3. Probability of obtaining a result in support of the hypothesis as a function of the number of names included in a study, under the assumption that the analyses conducted by the original authors are appropriate.

Figure 1

a) Names beginning with Tor



b) Names beginning with Edm

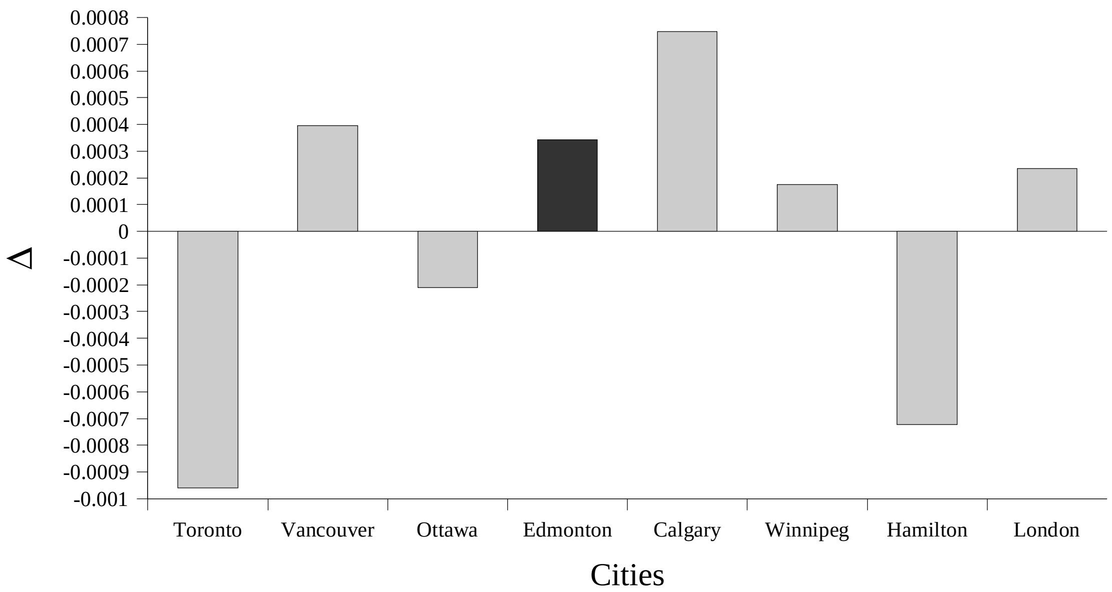
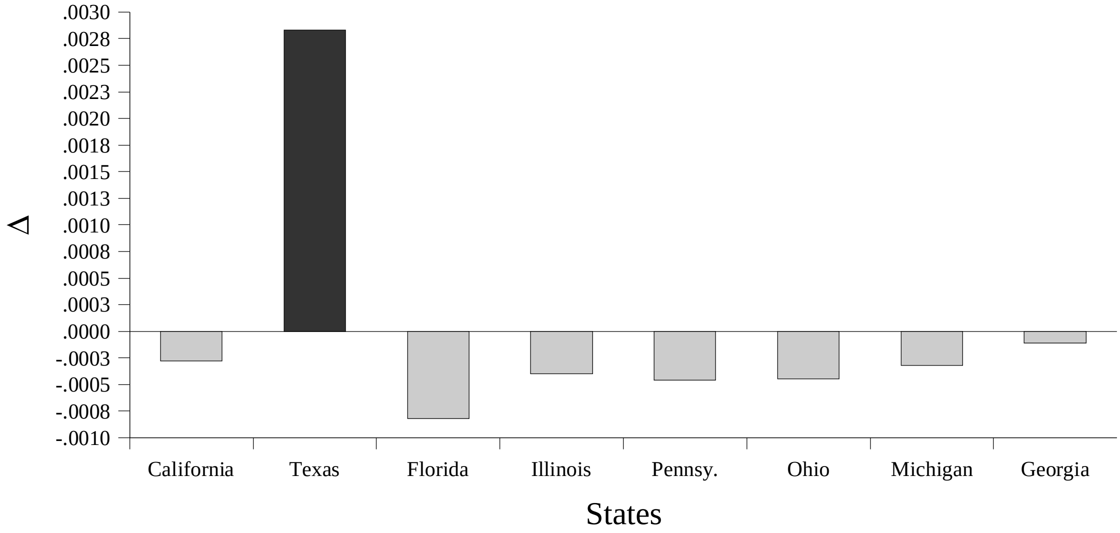


Figure 2

a) Names beginning with Texa



a) Names beginning with Illi

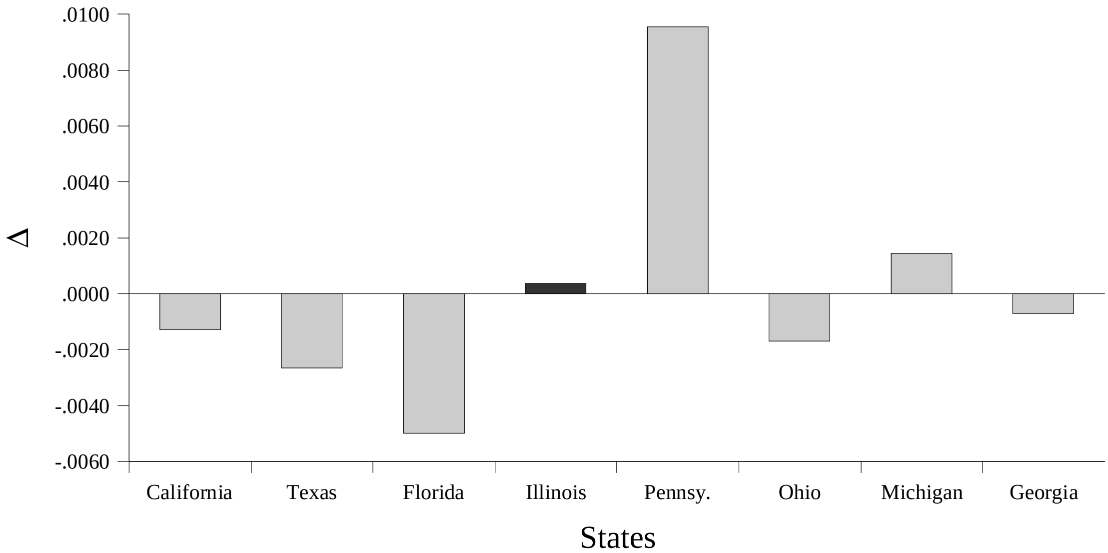
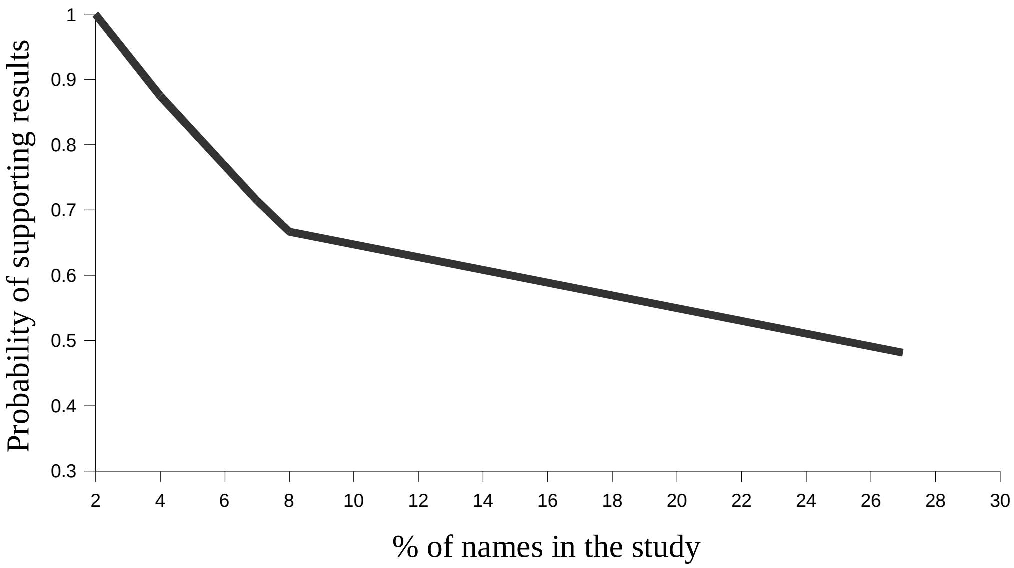


Figure 3


Probability of supporting results
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
% of names in the study